AD-A068 928 RICE UNIV HOUSTON TEX AERO-ASTRONAUTICS GROUP F/G 12/1
SEQUENTIAL CONJUGATE GRADIENT-RESTORATION ALGORITHM FOR OPTIMAL--ETC(U)
1978 A K WU, A MIELE AFOSR-76-3075
UNCLASSIFIED AAR-144 AFOSR-TR-79-0535 NL

1 OF 1
AD
A068928

END
DATE
FILMED
6 --79
DDC

AERO-ASTRONAUTICS REPORT NO. 144

SEQUENTIAL CONJUGATE GRADIENT-RESTORATION ALGORITHM FOR OPTIMAL CONTROL PROBLEMS WITH GENERAL BOUNDARY CONDITIONS

by

A. K. WU AND A. MIELE

RICE UNIVERSITY

1978

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 **REPORT DOCUMENTATION PAGE**

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 **AFOSR-TR- 79-0535**

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): 6 SEQUENTIAL CONJUGATE GRADIENT-RESTORATION ALGORITHM FOR OPTIMAL CONTROL PROBLEMS WITH GENERAL BOUNDARY CONDITIONS.

5. TYPE OF REPORT & PERIOD COVERED: Interim

6. PERFORMING ORG. REPORT NUMBER: 14 AAR-144

7. AUTHOR(s): 10 A.K./Wu [redacted] A./Miele

8. CONTRACT OR GRANT NUMBER(s): 15 AFOSR-76-3075

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Rice University
Department of Mechanical Engineering
Houston, Texas 77001

402 169

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
61102F
16 2304/A3
17 A3

11. CONTROLLING OFFICE NAME AND ADDRESS:
Air Force Office of Scientific Research/NM
Bolling AFB, DC 20332

12. REPORT DATE: 11 1978

13. NUMBER OF PAGES: 75

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

12 83p.

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

9 Aero-astronautics rept.,

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Optimal control, numerical methods, computing methods, gradient methods, gradient-restoration algorithms, sequential gradient-restoration algorithms, conjugate gradient-restoration algorithms, sequential conjugate gradient-restoration algorithms, general boundary conditions.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper considers the numerical solution of the problem of minimizing a functional subject to differential constraints and general boundary conditions.

The approach taken is a sequence of two-phase cycles, composed of a conjugate gradient phase and a restoration phase. The conjugate gradient phase involves one iteration and is designed to decrease the value of the functional, while the

to next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

402 169

20. ABSTRACT (Continued)

constraints are satisfied to first order. During this iteration, the first variation of the functional is minimized, subject to the linearized constraints and to a quadratic constraint imposed on the variations of the control, the parameter, and the missing components of the initial state. The restoration phase involves one or more iterations and is designed to force constraint satisfaction to a predetermined accuracy, while the norm squared of the variations of the control, the parameter, and the missing components of the initial state is minimized.

The algorithm has two principal properties: (i) it produces a sequence of feasible suboptimal solutions; the functions obtained at the end of each cycle satisfy the constraints to a predetermined accuracy; and (ii) the value of the functional at the end of any complete conjugate gradient-restoration cycle is smaller than the value of the same functional at the beginning of that cycle.

The sequential conjugate gradient-restoration algorithm presented here differs from previous work, in that it is not required that the state vector be given at the initial point. Instead, the initial conditions can be absolutely general. Since the present algorithm is capable of handling general final conditions, it is suitable for the solution of optimal control problems with general boundary conditions. Its importance lies in the fact that many optimal control problems involve initial conditions of the type considered here.

Nine numerical examples are presented in order to illustrate the performance of the algorithm. The numerical results show the feasibility as well as the convergence characteristics of the present algorithm.

LEVEL II

4

AERO-ASTRONAUTICS REPORT NO. 144

SEQUENTIAL CONJUGATE GRADIENT-RESTORATION ALGORITHM FOR OPTIMAL CONTROL PROBLEMS WITH GENERAL BOUNDARY CONDITIONS

by

A.K. WU and A. MIELE

RICE UNIVERSITY

1978

Sequential Conjugate Gradient-Restoration Algorithm for Optimal Control Problems with General Boundary Conditions[1]

by

A.K. WU[2] and A. MIELE[3]

Abstract. This paper considers the numerical solution of the problem of minimizing a functional I subject to differential constraints and general boundary conditions. It consists of finding the state x(t), the control u(t), and the parameter π so that the functional I is minimized while the constraints and the boundary conditions are satisfied to a predetermined accuracy.

---

[1] This research was supported by the Office of Scientific Research, Office of Aerospace Research, United States Air Force, Grant No. AF-AFOSR-76-3075. The authors are indebted to Dr. S. Gonzalez for analytical and computational assistance.

[2] Graduate Student, Department of Mechanical Engineering and Materials Science, Rice University, Houston, Texas.

[3] Professor of Astronautics and Mathematical Sciences, Rice University, Houston, Texas.

The approach taken is a sequence of two-phase cycles, composed of a conjugate gradient phase and a restoration phase. The conjugate gradient phase involves one iteration and is designed to decrease the value of the functional, while the constraints are satisfied to first order. During this iteration, the first variation of the functional is minimized, subject to the linearized constraints. The minimization is performed over the class of variations of the control, the parameter, and the missing components of the initial state which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase. The sequence of conjugate gradient phases generated by the algorithm is such that, for the special case of a quadratic functional subject to linear constraints, various orthogonality and conjugacy conditions hold. The restoration phase involves one or more iterations and is designed to force constraint satisfaction to a predetermined accuracy, while the norm squared of the variations of the control, the parameter, and the missing components of the initial state is minimized.

The principal property of the algorithm is that it produces a sequence of feasible suboptimal solutions: the functions obtained at the end of each cycle satisfy the constraints to a predetermined accuracy. Therefore, the values of the functional I corresponding to any two elements of the sequence are comparable.

The stepsize of the conjugate gradient phase is determined by a one-dimensional search on the augmented functional J, while the stepsize of the restoration phase is obtained by a one-dimensional search on the constraint error P. The conjugate gradient stepsize and the restoration stepsize are chosen so that the restoration phase preserves the descent property of the conjugate gradient phase. Therefore, the value of the functional I at the end of any complete conjugate gradient-restoration cycle is smaller than the value of the same functional at the beginning of that cycle. Of course, restarting the algorithm might be occasionally necessary.

The sequential conjugate gradient-restoration algorithm presented here differs from that of Refs. 3 and 4, in that it is not required that the state vector be given at the initial point. Instead, the initial conditions can be absolutely general. In analogy with Refs. 3 and 4, the present algorithm is capable of handling general final conditions; therefore, it is suitable for the solution of optimal control problems with general boundary conditions. Its importance lies in the fact that many optimal control problems involve initial conditions of the type considered here.

Nine numerical examples are presented in order to illustrate the performance of the algorithm. The numerical results show the feasibility as well as the convergence characteristics of the present algorithm.

Key Words. Optimal control, numerical methods, computing methods, gradient methods, gradient-restoration algorithms, sequential gradient-restoration algorithms, conjugate gradient-restoration algorithms, sequential conjugate gradient-restoration algorithms, general boundary conditions.

1. Introduction

Over the past several years, a successful family of algorithms for the solution of optimal control problems involving differential constraints and terminal constraints has been developed at Rice University by Miele and his associates (see Refs. 1-4). They are known as sequential gradient-restoration algorithms and have been designed for the solution of different classes of optimal control problems. Some of these algorithms are of the ordinary-gradient type (Refs. 1-2), while the rest are of the conjugate-gradient type (Refs. 2-4). All of them have shown to be robust and reliable; however, they all require the state vector to be given at the initial point. Owing to the fact that optimal control problems exist, which require satisfaction of more general boundary conditions, the task of extending the aforementioned family of algorithms must be undertaken, with these ideas in mind: (i) to retain the robustness, reliability, and convergence characteristics of the algorithms discussed in Refs. 1-4; (ii) to be able to handle all of the optimal control problems treated in Refs. 1-4; and (iii) to have the additional capability of handling optimal control problems with general boundary conditions.

In Ref. 5, an algorithm of the ordinary gradient type was developed, extending the work of Ref. 1 to problems with general boundary conditions. In this report, an algorithm of

the conjugate gradient type is developed, extending the work of Refs. 3-4 to problems with general boundary conditions.

Specifically, the following optimal control problem is considered: Minimize the functional I, which depends on the n-vector state x(t), the m-vector control u(t), and the p-vector parameter π. The state and the parameter are required to satisfy r scalar relations at the initial point and q scalar relations at the final point. Along the interval of integration, the state, the control, and the parameter are required to satisfy n scalar differential equations.

The approach taken is a sequence of two-phase cycles, composed of a conjugate gradient phase and a restoration phase. The conjugate gradient phase involves one iteration and is designed to decrease the value of the functional, while the constraints are satisfied to first order. During this iteration, the first variation of the functional is minimized, subject to the linearized constraints. The minimization is performed over the class of variations of the control, the parameter, and the missing components of the initial state which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase. The sequence of conjugate gradient phases generated by the algorithm is such that, for the special case of a quadratic functional subject to linear constraints, various orthogonality and conjugacy conditions hold. The restoration phase

involves one or more iterations and is designed to force constraint satisfaction to a predetermined accuracy, while the norm squared of the variations of the control, the parameter, and the missing components of the initial state is minimized.

The principal property of the algorithm is that it produces a sequence of feasible suboptimal solutions: the functions obtained at the end of each cycle satisfy the constraints to a predetermined accuracy. Therefore, the values of the functional I corresponding to any two elements of the sequence are comparable.

The stepsize of the conjugate gradient phase is determined by a one-dimensional search on the augmented functional J, while the stepsize of the restoration phase is obtained by a one-dimensional search on the constraint error P. The conjugate gradient stepsize and the restoration stepsize are chosen so that the restoration phase preserves the descent property of the conjugate gradient phase. Therefore, the value of the functional I at the end of any complete conjugate gradient-restoration cycle is smaller than the value of the same functional at the beginning of that cycle. Of course, restarting the algorithm might be occasionally necessary.

A time normalization is used in order to simplify the numerical computations. Specifically, the actual time θ is replaced by the normalized time $t = \theta/\tau$, which is defined in

such a way that the initial time is $t = 0$ and the final time is $t = 1$. The actual final time τ, if it is free, is regarded as a component of the vector parameter π to be optimized. In this way, an optimal control problem with variable final time is converted into an optimal control problem with fixed final time.

2. Statement of the Problem

2.1. Notation. Let t denote the independent variable, and let $x(t)$, $u(t)$, π denote the dependent variables. The time t is a scalar, the state $x(t)$ is an n-vector, the control $u(t)$ is an m-vector, and the parameter π is a p-vector.[4]

The state $x(t)$ is partitioned into vectors $y(t)$ and $z(t)$, defined as follows: $y(t)$ is an a-vector including those components of the state that are prescribed at the initial point, and $z(t)$ is a b-vector including those components of the state that are not prescribed at the initial point. Clearly, $a + b = n$.

2.2. Optimization Problem. With the above notations, the optimization problem can be stated as follows. Minimize the functional

$$I = \int_0^1 f(x,u,\pi,t)dt + [h(z,\pi)]_0 + [g(x,\pi)]_1 , \qquad (1)$$

with respect to the state $x(t)$, the control $u(t)$, and the parameter π which satisfy the differential constraints

$$\dot{x} - \phi(x,u,\pi,t) = 0 , \qquad 0 \leq t \leq 1, \qquad (2)$$

and the boundary conditions

[4]All vectors are column vectors.

$$y(0) = \text{given}, \tag{3}$$

$$[\omega(z,\pi)]_0 = 0\,, \tag{4}$$

$$[\psi(x,\pi)]_1 = 0\,. \tag{5}$$

In the above equations, the quantities I,f,h,g are scalar, the function ϕ is an n-vector, the function ω is a c-vector, and the function ψ is a q-vector. The number of initial conditions $r = a + c$ and the number of final conditions q must satisfy the following relation:

$$r + q \le n_0 + n_1 + p_* \le 2n + p\,. \tag{6}$$

In Ineq. (6), the symbol $p_* \le p$ denotes the number of components of the parameter π present in the boundary conditions; the symbol $n_0 \le n$ denotes the number of state variables present in the initial conditions; and the symbol $n_1 \le n$ denotes the number of state variables present in the final conditions.

2.3. First-Order Conditions. From calculus of variations, it can be seen that the previous problem is one of the Bolza type, and it can be recast as that of minimizing the augmented functional

$$J = I + L\,, \tag{7-1}$$

subject to (2)-(5). In compact notation, the functional I

can be rewritten as

$$I = \int_0^1 f\,dt + (h)_0 + (g)_1\,, \tag{7-2}$$

and the functional L is defined as

$$L = \int_0^1 \lambda^T(\dot{x} - \phi)\,dt + (\sigma^T\omega)_0 + (\mu^T\psi)_1\,, \tag{7-3}$$

where the n-vector $\lambda(t)$ is a variable Lagrange multiplier, the c-vector σ is a constant Lagrange multiplier, and the q-vector μ is a constant Lagrange multiplier. After integrating by parts the term $\lambda^T\dot{x}$, the functional (7-3) can be rewritten as

$$L = \int_0^1 (-\dot{\lambda}^T x - \lambda^T\phi)\,dt + (-\lambda^T x + \sigma^T\omega)_0 + (\lambda^T x + \mu^T\psi)_1\,. \tag{7-4}$$

The functions $x(t)$, $u(t)$, π and the multipliers $\lambda(t)$, σ, μ solving the previous problem must satisfy the feasibility equations (2)-(5) and the following optimality conditions:

$$\dot{\lambda} - f_x + \phi_x\lambda = 0, \qquad 0 \le t \le 1, \tag{8}$$

$$f_u - \phi_u\lambda = 0, \qquad 0 \le t \le 1, \tag{9}$$

$$\int_0^1 (f_\pi - \phi_\pi\lambda)\,dt + (h_\pi + \omega_\pi\sigma)_0 + (g_\pi + \psi_\pi\mu)_1 = 0, \tag{10}$$

$$(-\zeta + h_z + \omega_z \sigma)_0 = 0, \tag{11}$$

$$(\lambda + g_x + \psi_x \mu)_1 = 0 . \tag{12}$$

2.4. Remark. Just as the state vector x(t) is partitioned into an a-vector y(t) and a b-vector z(t), the multiplier vector $\lambda(t)$ is partitioned into a a-vector $\eta(t)$ and a b-vector $\zeta(t)$, having the following meaning: $\eta(t)$ is associated with y(t), and $\zeta(t)$ is associated with z(t). With reference to Eq. (11), $\zeta(0)$ denotes the portion of the initial Lagrange multiplier $\lambda(0)$ which is associated with z(0), the portion of the initial state vector x(0) which is not prescribed.

2.5. Approximate Methods. Since in general the differential system (2)-(5) and (8)-(12) is nonlinear, approximate methods are employed to find a solution iteratively. In this connection, let the norm squared of a vector ν be defined by

$$N(\nu) = \nu^T \nu . \tag{13}$$

Then, the constraint error P can be written as[5]

$$P = \int_0^1 N(\dot{x} - \phi)\,dt + N(\omega)_0 + N(\psi)_1 , \tag{14}$$

[5]In Eq. (14), it is assumed that the initial condition (3) is satisfied.

and the error in the optimality conditions Q is given by

$$Q = \int_0^1 N(\dot{\lambda} - f_x + \phi_x \lambda)dt + \int_0^1 N(f_u - \phi_u \lambda)dt$$

$$+ N\left[\int_0^1 (f_\pi - \phi_\pi \lambda)dt + (h_\pi + \omega_\pi \sigma)_0 + (g_\pi + \psi_\pi \mu)_1\right]$$

$$+ N(-\zeta + h_z + \omega_z \sigma)_0 + N(\lambda + g_x + \psi_x \mu)_1 . \tag{15}$$

For the exact optimal solution, one must have

$$P = 0, \qquad Q = 0. \tag{16}$$

For an approximation to the optimal solution, the following relations are to be satisfied:

$$P \leq \varepsilon_1 , \qquad Q \leq \varepsilon_2 , \tag{17}$$

where ε_1 and ε_2 are small, preselected numbers.

3. Description of the Algorithm

The technique employed is characterized by a sequence of two-phase cycles, composed of a conjugate gradient phase and a restoration phase. These phases are described below.

The conjugate gradient phase is started only when the constraint error P satisfies Ineq. (17-1). It involves a single iteration, which is designed to decrease the value of the functional I or the augmented functional J, while the constraints are satisfied to first order. During this iteration, the first variation of the functional I is minimized, subject to the linearized constraints. The minimization is performed over the class of variations of the control $u(t)$ and the parameters π and $z(0)$ which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase.[6]

The restoration phase is started only when the constraint error P violates Ineq. (17-1). The restoration phase involves one or more iterations. In each restorative iteration, the objective is to reduce the constraint error P, while the

[6] The sequence of conjugate gradient phases generated by the algorithm is such that, for the special case of a quadratic functional subject to linear constraints, various orthogonality and conjugacy conditions hold (see Section 6).

constraints are satisfied to first order and the norm squared of the variations of the control $u(t)$ and the parameters π and $z(0)$ is minimized. The restoration phase is terminated whenever Ineq. (17-1) is satisfied.

A complete conjugate gradient-restoration cycle is designed so that the value of the functional I decreases, while the constraints are satisfied to the accuracy (17-1) both at the beginning and at the end of the cycle. Finally, the algorithm as a whole is terminated whenever Ineqs. (17) are satisfied simultaneously.

3.1. Remark. During each conjugate gradient iteration or restorative iteration, use of nonlinear equations must be avoided. *Therefore*, the exact feasibility equations (2)-(5) are replaced by their corresponding linearized approximations. These linearized approximations do not include forcing terms in the conjugate gradient phase, while they do include forcing terms in the restoration phase.

3.2. Notation. For any iteration of the conjugate gradient phase or the restoration phase, the following terminology is adopted: $x(t)$, $u(t)$, π denote the nominal functions; $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ denote the varied functions; and $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ denote the displacements leading from the nominal functions to the varied functions. These quantities satisfy the relations

$$\tilde{x}(t) = x(t) + \Delta x(t), \qquad \tilde{u}(t) = u(t) + \Delta u(t), \qquad \tilde{\pi} = \pi + \Delta\pi . \tag{18}$$

Because the vector x is partitioned into y and z, Eq. (18-1) implies that

$$\tilde{y}(t) = y(t) + \Delta y(t), \qquad \tilde{z}(t) = z(t) + \Delta z(t). \tag{19}$$

Let α be a positive number representing the stepsize (either the gradient stepsize or the restoration stepsize). Then, we define the displacements per unit stepsize as follows:

$$A(t) = \Delta x(t)/\alpha, \qquad B(t) = \Delta u(t)/\alpha, \qquad C = \Delta\pi/\alpha. \tag{20}$$

The vector A is partitioned into vectors D and E associated with y and z, respectively. Therefore, Eq. (20-1) implies that

$$D(t) = \Delta y(t)/\alpha, \qquad E(t) = \Delta z(t)/\alpha. \tag{21}$$

Upon combining (18)-(19) with (20)-(21), we see that

$$\tilde{x}(t) = x(t) + \alpha A(t), \qquad \tilde{u}(t) = u(t) + \alpha B(t), \qquad \tilde{\pi} = \pi + \alpha C, \tag{22}$$

$$\tilde{y}(t) = y(t) + \alpha D(t), \qquad \tilde{z}(t) = z(t) + \alpha E(t). \tag{23}$$

3.3. Desired Properties. The functions $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ must be determined so as to produce some desirable effect at every iteration, namely, the decrease of the functionals I, and/or J, and/or P. Thus, the following descent properties are required:

$$\tilde{I} < I, \quad \text{and/or} \quad \tilde{J} < J, \quad \text{and/or} \quad \tilde{P} < P, \tag{24}$$

where I, J, P are associated with the nominal functions and $\tilde{I}$, $\tilde{J}$, $\tilde{P}$ are associated with the varied functions. In turn, the functions A(t), B(t), C are chosen so that

$$\delta I < 0, \quad \text{and/or} \quad \delta J < 0, \quad \text{and/or} \quad \delta P < 0, \tag{25}$$

where the symbol $\delta(\ldots)$ denotes the first variation. Then, by choosing the stepsize α sufficiently small, the satisfaction of relations (24) is guaranteed. Ineqs. (24-1), (24-2) and (25-1), (25-2) characterize the conjugate gradient phase, while Ineqs. (24-3) and (25-3) characterize the restoration phase.

3.4. First Variations. Next, we give the expressions for the first variations of the functionals I, J, P; after simple manipulations, omitted for the sake of brevity, they take the form[7,8]

$$\delta I/\alpha = \int_0^1 (f_x^T A + f_u^T B + f_\pi^T C)\,dt + (h_z^T E + h_\pi^T C)_0 + (g_x^T A + g_\pi^T C)_1, \tag{26}$$

[7]Implicit in Eqs. (26)-(28) is the assumption $D(0) = 0$.

[8]The first variation of the augmented functional J is computed by varying the functions x(t), u(t), π, while holding the multipliers $\lambda(t), \sigma, \mu$ unchanged.

and

$$\delta J/\alpha = \int_0^1 (-\dot{\lambda} + f_x - \phi_x\lambda)^T A dt + \int_0^1 (f_u - \phi_u\lambda)^T B dt$$

$$+ \left[\int_0^1 (f_\pi - \phi_\pi\lambda) dt + (h_\pi + \omega_\pi\sigma)_0 + (g_\pi + \psi_\pi\mu)_1\right]^T C$$

$$+ \left[(-\zeta + h_z + \omega_z\sigma)^T E\right]_0 + \left[(\lambda + g_x + \psi_x\mu)^T A\right]_1, \tag{27}$$

and

$$\delta P/2\alpha = \int_0^1 (\dot{x} - \phi)^T (\dot{A} - \phi_x^T A - \phi_u^T B - \phi_\pi^T C) dt$$

$$+ \left[\omega^T(\omega_z^T E + \omega_\pi^T C)\right]_0 + \left[\psi^T(\psi_x^T A + \psi_\pi^T C)\right]_1. \tag{28}$$

3.5. <u>Remark</u>. For the purposes of this report, Eqs. (26)-(28) must be completed by one of the following relations:

$$K/\alpha^2 = \int_0^1 B^T B dt + C^T C + (E^T E)_0, \tag{29}$$

or

$$K/\alpha^2 = \int_0^1 (B-\gamma\hat{B})^T (B-\gamma\hat{B}) dt + (C-\gamma\hat{C})^T (C-\gamma\hat{C}) + \left[(E-\gamma\hat{E})^T (E-\gamma\hat{E})\right]_0, \tag{30}$$

which constitute a measure of the overall change of the control $u(t)$ and the parameters π and $z(0)$. Equation (29) is employed in the restoration phase, and Eq. (30) is employed in the conjugate gradient phase. In Eq. (30), the functions $A(t)$, $B(t)$, C pertain to the present conjugate gradient phase, and the functions $\hat{A}(t)$, $\hat{B}(t)$, $\hat{C}$ pertain to the previous conjugate gradient phase. The symbol γ denotes a scalar, nonnegative quantity, called the directional coefficient. Its specification is discussed in Section 6.

4. Restoration Phase

4.1. Linearized Equations. Let $x(t)$, $u(t)$, π denote nominal functions satisfying (3), but not necessarily (2), (4), (5). To first order, the perturbations per unit stepsize $A(t)$, $B(t)$, C must satisfy the linearized constraint equations

$$\dot{A} - \phi_x^T A - \phi_u^T B - \phi_\pi^T C + (\dot{x} - \phi) = 0, \qquad 0 \le t \le 1, \tag{31}$$

$$D(0) = 0, \tag{32}$$

$$(\omega_z^T E + \omega_\pi^T C + \omega)_0 = 0\ , \tag{33}$$

$$(\psi_x^T A + \psi_\pi^T C + \psi)_1 = 0\ . \tag{34}$$

4.2. Descent Property. The linearized equations (31)-(34) admit an infinite number of solutions, each of which is characterized by a descent property in the constraint error P. This descent property can be seen by combining (28) with (31)-(34): the first variation of P becomes

$$\delta P = -2\alpha P\ . \tag{35}$$

Since $P > 0$, Eq. (35) shows that $\delta P < 0$; hence, for α sufficiently small, a decrease in the constraint error P is guaranteed.

4.3. Special Variations. Among the infinite number of solutions of Eqs. (31)-(34), the one that minimizes the

functional (29) is selected. Thus, we seek the minimum of the quadratic functional (29), with respect to the functions A(t), B(t), C which satisfy the linearized constraints (31)-(34).

By applying standard techniques of optimal control theory or calculus of variations, the following optimality conditions are obtained:

$$B = \phi_u \lambda, \qquad 0 \leq t \leq 1, \tag{36}$$

$$C = \int_0^1 \phi_\pi \lambda dt - (\omega_\pi \sigma)_0 - (\psi_\pi \mu)_1, \tag{37}$$

$$E(0) = (\zeta - \omega_z \sigma)_0, \tag{38}$$

$$\dot{\lambda} + \phi_x \lambda = 0, \qquad 0 \leq t \leq 1, \tag{39}$$

$$(\lambda + \psi_x \mu)_1 = 0. \tag{40}$$

Summarizing, we seek functions A(t), B(t), C and multipliers $\lambda(t)$, σ, μ which satisfy the linearized constraints (31)-(34) and the optimality conditions (36)-(40).

4.4. Linear, Two-Point Boundary-Value Problem. The technique used to solve the LTP-BVP (31)-(34) and (36)-(40) is a backward-forward integration scheme in combination with the method of particular solutions (Refs. 6-8). The technique requires the execution of $q+1$ independent sweeps of the

differential system (31)-(34) and (36)-(40), each characterized by a different value of the multiplier μ.

The generic sweep is started by assigning particular values to the components of μ; then, the multiplier $\lambda(1)$ is obtained from (40). Next, Eq. (39) is integrated backward to obtain the function $\lambda(t)$, and Eq. (36) is employed to obtain $B(t)$. With $\lambda(0)$ known, $\zeta(0)$ is known. Therefore, Eqs. (33), (37), (38) constitute a system of $b + c + p$ linear relations in which the unknowns are the $b + c + p$ components of the vectors $E(0)$, σ, C. For this system to have a unique solution, the following disequation must hold:[9]

$$\det\left[\omega_z^T\omega_z + \omega_\pi^T\omega_\pi\right]_0 \neq 0. \tag{41}$$

With $E(0)$ known and because of (32), the vector $A(0)$ is known. Then, $A(t)$ is obtained by forward integration of (31). In this way, the sweep is completed: for the arbitrary value assigned to μ, it leads to the satisfaction of all of the equations of the system (31)-(34) and (36)-(40), except Eq. (34).

---

[9]Disequation (41) is obtained from (33), (37), (38) after elimination of $E(0)$ and C. The resulting linear equation in σ admits a unique solution providing (41) is satisfied.

In order to satisfy Eq. (34) and because the system (31)-(34) and (36)-(40) is nonhomogeneous, $q+1$ independent sweeps must be executed employing $q+1$ different multiplier vectors μ_i, $i=1,\ldots,q+1$. The first q sweeps are performed by choosing the vectors $\mu_1,\ldots\ldots,\mu_q$ to be the columns of the identity matrix of order q. The last sweep is executed by choosing μ_{q+1} to be the null vector. As a result, one generates the functions and multipliers

$$A_i(t),\ B_i(t),\ C_i,\ \lambda_i(t),\ \sigma_i,\ \mu_i,\ i=1,\ldots\ldots,q+1. \tag{42}$$

Now, we introduce the $q+1$ undetermined, scalar constants k_i and form the linear combinations

$$A(t)=\Sigma k_iA_i(t), \qquad B(t)=\Sigma k_iB_i(t), \qquad C=\Sigma k_iC_i, \tag{43}$$

$$\lambda(t)=\Sigma k_i\lambda_i(t), \qquad \sigma=\Sigma k_i\sigma_i, \qquad \mu=\Sigma k_i\mu_i, \tag{44}$$

where the summations are taken over the index i. The $q+1$ coefficients k_i are obtained by forcing the linear combinations (43) to satisfy Eq. (34), together with the normalization condition (Ref. 6)

$$\Sigma k_i=1\ . \tag{45}$$

Once the constants k_i are known, the solution of the LTP-BVP (31)-(34) and (36)-(40) is given by (43)-(44).

4.5. <u>Restoration Stepsize</u>. With the functions $A(t)$, $B(t)$, C known, the one-parameter family of varied functions (22) can be formed. For this one-parameter family, the constraint error (14) becomes a function of the form

$$\tilde{P} = \tilde{P}(\alpha) . \tag{46}$$

Then, the stepsize α must be selected so that the following relations are satisfied:

$$\tilde{P}(\alpha) < \tilde{P}(0), \qquad \tilde{\tau}(\alpha) > 0. \tag{47}$$

Satisfaction of Ineq. (47-1) is possible because of the descent property of the restoration phase. Ineq. (47-2) is required for problems with free final time.

In order to achieve satisfaction of (47), a bisection process is applied to the restoration stepsize α, starting from the reference stepsize $\alpha_o = 1$. This reference stepsize has the property of yielding one-step restoration for the case where the constraints (2)-(5) are linear.

4.6. <u>Iterative Procedure for the Restoration Phase</u>. The descent property (35) of the restoration phase guarantees satisfaction of Ineq. (47-1) at the end of any iteration, but not satisfaction of Ineq. (17-1). Therefore, the restoration algorithm must be employed iteratively until Ineq. (17-1) is satisfied. At this point, the restoration phase is terminated.

5. Conjugate Gradient Phase: General Case

5.1. Linearized Equations. Suppose that nominal functions $x(t)$, $u(t)$, π are available, which satisfy (2)-(5). To first order, the perturbations per unit stepsize $A(t)$, $B(t)$, C must satisfy the linearized constraint equations

$$\dot{A} - \phi_x^T A - \phi_u^T B - \phi_\pi^T C = 0, \qquad 0 \leq t \leq 1, \tag{48}$$

$$D(0) = 0, \tag{49}$$

$$(\omega_z^T E + \omega_\pi^T C)_0 = 0, \tag{50}$$

$$(\psi_x^T A + \psi_\pi^T C)_1 = 0. \tag{51}$$

5.2. Special Variations. Among the infinite number of solutions of Eqs. (48)-(51), the one that minimizes the functional (26) is selected. Thus, we seek the minimum of the linear functional (26), with respect to the functions $A(t)$, $B(t)$, C which satisfy the linearized constraints (48)-(51) and the quadratic isoperimetric constraint (30).

By applying standard techniques of optimal control theory or calculus of variations, the following optimality conditions are obtained:[10]

---

[10] In Eqs. (52)-(56), the multiplier ν associated with the isoperimetric constraint (30) is set at the level $2\nu = 1$.

$$B = \gamma\hat{B} - (f_u - \phi_u\lambda), \qquad 0 \leq t \leq 1, \tag{52}$$

$$C = \gamma\hat{C} - \left[\int_0^1 (f_\pi - \phi_\pi\lambda)dt + (h_\pi + \omega_\pi\sigma)_0 + (g_\pi + \psi_\pi\mu)_1\right], \tag{53}$$

$$E(0) = \gamma\hat{E}(0) - (-\zeta + h_z + \omega_z\sigma)_0, \tag{54}$$

$$\dot{\lambda} - f_x + \phi_x\lambda = 0, \qquad 0 \leq t \leq 1, \tag{55}$$

$$(\lambda + g_x + \psi_x\mu)_1 = 0. \tag{56}$$

Summarizing, for a given value of the directional coefficient γ, we seek functions $A(t)$, $B(t)$, C and multipliers $\lambda(t)$, σ, μ which satisfy the linearized constraints (48)-(51) and the optimality conditions (52)-(56).

5.3. Isoperimetric Constraint. In the light of (52)-(56), the error in the optimality conditions (15) reduces to

$$Q = \int_0^1 (B-\gamma\hat{B})^T(B-\gamma\hat{B})dt + (C-\gamma\hat{C})^T(C-\gamma\hat{C}) + [(E-\gamma\hat{E})^T(E-\gamma\hat{E})]_0 . \tag{57}$$

Consequently, the following relation ties the isoperimetric constant, the stepsize, and the error in the optimality conditions:

$$K = \alpha^2 Q . \tag{58}$$

Clearly, to assign values to the isoperimetric constant is the same as assigning values to the stepsize. However, there is no clear-cut way of determining a priori convenient values for the isoperimetric constant. Therefore, the implementation of the conjugate gradient algorithm becomes simpler if one avoids evaluating α in terms of K through (58) and assigns values to α directly.

5.4. Descent Property. When the variations defined by (48)-(56) are employed, the first variation of the augmented functional (27) becomes

$$\delta J = -\alpha(Q + \gamma Z), \tag{59}$$

where Q is given by (57) and

$$Z = \int_0^1 (B - \gamma\hat{B})^T \hat{B} dt + (C - \gamma\hat{C})^T \hat{C} + [(E - \gamma\hat{E})^T \hat{E}]_0 . \tag{60}$$

For the first iteration of the conjugate gradient phase, one sets

$$\gamma = 0 , \tag{61}$$

with the implication that

$$\delta J = -\alpha Q . \tag{62}$$

Since $Q > 0$, Eq. (62) shows that $\delta J < 0$. Hence, for α sufficiently small, it is guaranteed that the augmented functional J decreases.

For subsequent iterations, one sets $\gamma \neq 0$. More specifically, the directional coefficient must be such that

$$\gamma > 0 , \tag{63}$$

and its proper value is discussed in Section 6. At any rate, Eq. (59) shows that $\delta J < 0$ providing

$$Q + \gamma Z > 0 . \tag{64}$$

Hence, for α sufficiently small, it is guaranteed that the augmented functional J decreases as long as Ineq. (64) is satisfied. If Ineq. (64) is violated, the descent property on J no longer holds, and the conjugate gradient phase must be restarted by resetting the directional coefficient γ at the level (61).

5.5. Linear, Two-Point Boundary-Value Problem. For a given value of the directional coefficient γ, the technique used to solve the LTP-BVP (48)-(56), associated with the conjugate gradient phase, is analogous to that described for the restoration phase (see Section 4.4); hence, it is not repeated, for the sake of brevity.

5.6. General Solution. Next, assume that two particular values are given to the directional coefficient γ, for instance,

$$\gamma_* = 0 \qquad \text{and} \qquad \gamma_{**} = 1. \tag{65}$$

Denote by

$$A_*(t),\ B_*(t),\ C_*,\ \lambda_*(t),\ \sigma_*,\ \mu_* \tag{66}$$

and

$$A_{**}(t),\ B_{**}(t),\ C_{**},\ \lambda_{**}(t),\ \sigma_{**},\ \mu_{**} \tag{67}$$

the particular solutions of the LTP-BVP (48)-(56) corresponding to (65-1) and (65-2), respectively. Simple manipulations, omitted for the sake of brevity, show that the general solution of (48)-(56), valid for any value of the directional coefficient γ, can be written as

$$A(t) = A_*(t) + \gamma[A_{**}(t) - A_*(t)]\ , \tag{68-1}$$

$$B(t) = B_*(t) + \gamma[B_{**}(t) - B_*(t)]\ , \tag{68-2}$$

$$C = C_* + \gamma(C_{**} - C_*)\ , \tag{68-3}$$

and

$$\lambda(t) = \lambda_*(t) + \gamma[\lambda_{**}(t) - \lambda_*(t)]\ , \tag{69-1}$$

$$\sigma = \sigma_* + \gamma(\sigma_{**} - \sigma_*) \ , \tag{69-2}$$

$$\mu = \mu_* + \gamma(\mu_{**} - \mu_*) \ . \tag{69-3}$$

As a conclusion, the general solution of (48)-(56) requires that two sets of $q+1$ sweeps be executed, one leading to the particular solution (66) and one leading to the particular solution (67).

5.7. Stepsize and Directional Coefficient. With the functions A(t), B(t), C known, the following two-parameter family of varied functions can be formed:

$$\tilde{x}(t) = x(t) + \alpha\left\{A_*(t) + \gamma[A_{**}(t) - A_*(t)]\right\} \ , \tag{70-1}$$

$$\tilde{u}(t) = u(t) + \alpha\left\{B_*(t) + \gamma[B_{**}(t) - B_*(t)]\right\} \ , \tag{70-2}$$

$$\tilde{\pi} = \pi + \alpha[C_* + \gamma(C_{**} - C_*)] \ . \tag{70-3}$$

On the other hand, the multipliers $\lambda(t)$, σ, μ form the one-parameter family (69). Upon using (69) and (70), we see that the augmented functional (7) takes the form

$$\tilde{J} = \tilde{J}(\alpha, \gamma) \ . \tag{71}$$

Therefore, the optimum values of α and γ satisfy the relations

$$\tilde{J}_\alpha(\alpha, \gamma) = 0, \qquad \tilde{J}_\gamma(\alpha, \gamma) = 0 \ . \tag{72}$$

Since the simultaneous determination of α and γ might be be expensive computationally, we proceed in a different way. First, we determine an approximate value of the directional coefficient γ, based on the consideration of the linear-quadratic model (Section 6). Once γ is known, the two-parameter family (71) reduces to the one-parameter family

$$\tilde{J} = \tilde{J}(\alpha) \ . \tag{73}$$

Then, the optimum stepsize α satisfies the relation

$$\tilde{J}_{\alpha}(\alpha) = 0 \ , \tag{74}$$

whose numerical solution can be obtained using quadratic interpolation or cubic interpolation (Ref. 9).

6. Conjugate Gradient Phase: Linear-Quadratic Case

In the previous section, we analyzed the conjugate gradient phase in general, regardless of the analytical form of the functional (1) and the constraints (2)-(5). In this section, we consider the linear-quadratic case, that is, the case where the functional (1) is quadratic and the constraints (2)-(5) are linear.

6.1. General Solution. Under the assumption of linear constraints, it can be verified that the particular solutions (66) and (67) satisfy the relations

$$A_{**}(t)-A_{*}(t)=\hat{A}(t), \qquad B_{**}(t)-B_{*}(t)=\hat{B}(t), \qquad C_{**}-C_{*}=\hat{C}, \tag{75}$$

$$\lambda_{**}(t)-\lambda_{*}(t)=0, \qquad \sigma_{**}-\sigma_{*}=0, \qquad \mu_{**}-\mu_{*}=0. \tag{76}$$

As a consequence, Eqs. (68)-(69) reduce to

$$A(t)=A_{*}(t)+\gamma\hat{A}(t), \qquad B(t)=B_{*}(t)+\gamma\hat{B}(t), \qquad C=C_{*}+\gamma\hat{C}, \tag{77}$$

$$\lambda(t)=\lambda_{*}(t), \qquad \sigma=\sigma_{*}, \qquad \mu=\mu_{*}. \tag{78}$$

This means that the general solution of the LTP-BVP (48)-(56) can be obtained by executing only one set of $q+1$ sweeps, namely, the set of sweeps leading to the solution (66). By the way, this is the solution corresponding to (65-1), namely, the solution associated with the ordinary gradient phase of Ref. 5.

6.2. Local Orthogonality Conditions. Under the assumption of linear constraints, the two-parameter family (70) simplifies to

$$\tilde{x}(t) = x(t) + \alpha[A_*(t) + \gamma\hat{A}(t)] \,, \tag{79-1}$$

$$\tilde{u}(t) = u(t) + \alpha[B_*(t) + \gamma\hat{B}(t)] \,, \tag{79-2}$$

$$\tilde{\pi} = \pi + \alpha(C_* + \gamma\hat{C}) \,. \tag{79-3}$$

On the other hand, the multipliers $\lambda(t)$, σ, μ are given by Eqs. (78). Upon using (78) and (79), we see that the augmented functional (7) still takes the form (71). Hence, the optimum values of α and γ still satisfy the relations (72).

After laborious manipulations, omitted for the sake of brevity, Eqs. (72) lead to the following local orthogonality conditions:

$$\int_0^1 \tilde{B}_*^T B\,dt + \tilde{C}_*^T C + (\tilde{E}_*^T E)_0 = 0, \tag{80-1}$$

$$\int_0^1 \tilde{B}_*^T \hat{B}\,dt + \tilde{C}_*^T \hat{C} + (\tilde{E}_*^T \hat{E})_0 = 0, \tag{30-2}$$

with the implication that

$$\int_0^1 \tilde{B}_*^T B_*\,dt + \tilde{C}_*^T C_* + (\tilde{E}_*^T E_*)_0 = 0 \,. \tag{80-3}$$

Here, the adjective local is employed to mean that Eqs. (80) involve vectors B(t), C, E(0) which are solutions of (48)-(56) computed for the present iteration and the previous iteration; they also involve vectors $B_*(t)$, C_*, $E_*(0)$ which are solutions of (48)-(56) for $\gamma = 0$ computed for the present iteration and the next iteration.

6.3. Local Conjugacy Condition. Let w(t) and M(t) denote the vectors

$$w(t) = \begin{bmatrix} x(t) \\ u(t) \\ \pi \end{bmatrix}, \qquad M(t) = \begin{bmatrix} A(t) \\ B(t) \\ C \end{bmatrix} . \tag{81}$$

Let f_{ww}, g_{ww}, h_{ww} denote the Hessian matrices of the functions f,g,h with respect to the vector w. With this notation, and under the assumption of linear constraints and quadratic functional, Eqs. (72) lead to the following local conjugacy condition:

$$\int_0^1 M^T f_{ww} \hat{M} dt + (M^T h_{ww} \hat{M})_0 + (M^T g_{ww} \hat{M})_1 = 0 . \tag{82}$$

Here, the adjective local is employed to mean that Eq. (82) involves vectors M(t), that is, vectors A(t), B(t), C,which are solutions of (48)-(56) computed for the present iteration and the previous iteration.

6.4. Stepsize and Directional Coefficient. After observing that

$$M(t) = M_*(t) + \gamma \hat{M}(t) , \tag{83}$$

and after accounting for Eqs. (80) and (82), it can be shown that the optimal values of γ and α are given by

$$\gamma = Q/\hat{Q} \quad , \quad \alpha = Q/R , \tag{84}$$

where

$$Q = \int_0^1 B_*^T B_* dt + C_*^T C_* + (E_*^T E_*)_0 , \tag{85-1}$$

$$\hat{Q} = \int_0^1 \hat{B}_*^T \hat{B}_* dt + \hat{C}_*^T \hat{C}_* + (\hat{E}_*^T \hat{E}_*)_0 , \tag{85-2}$$

$$R = \int_0^1 M^T f_{ww} M dt + (M^T h_{ww} M)_0 + (M^T g_{ww} M)_1 . \tag{85-3}$$

Clearly, the optimal directional coefficient γ is the ratio of the error in the optimality conditions Q for the present conjugate gradient iteration to the error in the optimality conditions $\hat{Q}$ for the previous conjugate gradient iteration. These quantities are known, since they involve vectors $B_*(t)$, C_*, $E_*(0)$ which are solutions of (48)-(56) for

$\gamma = 0$ computed for the present iteration and the previous iteration.

With γ known, the vector $M(t)$ is computed with (83) and the scalar quantity R is computed with (85-3). Then, the optimal stepsize α is evaluated with (84-2). Clearly, the optimal stepsize α is the ratio of the error in the optimality conditions Q to the scalar quantity R, which constitutes a measure of the curvature of the functional (1). Both Q and R are computed employing quantities pertaining to the present conjugate gradient iteration.

6.5. Descent Property. Under the assumption of linear constraints, Eq. (60) simplifies to

$$Z = \int_0^1 B_*^T \hat{B} dt + C_*^T \hat{C} + (E_*^T \hat{E})_0 . \tag{86}$$

Because of the local orthogonality condition (80-1) written for the previous iteration, Eq. (86) yields

$$Z = 0 . \tag{87}$$

As a consequence, Eq. (59) reduces to

$$\delta J = -\alpha Q , \tag{88}$$

where the error in the optimality conditions Q is given by Eq. (85-1). Equation (88) holds for any conjugate gradient

iteration and shows that, since $Q > 0$, we have $\delta J < 0$. Hence, for α sufficiently small, it is guaranteed that the augmented functional J decreases. In conclusion, for the linear-quadratic case, the restart procedure mentioned in Section 5.4 never occurs. This means that the directional coefficient γ is set at the level (61) only for the first conjugate gradient iteration.

6.6. <u>General Orthogonality and Conjugacy Conditions</u>. Now, assume that the algorithm described by Eqs. (48)-(56) and (79) is employed, starting with some feasible nominal functions. Further, assume that the directional coefficient γ is set at the level (61) for the first conjugate gradient iteration and at the level (84-1) for any subsequent conjugate gradient iteration. Under these assumptions and for the linear-quadratic case, one can generalize the local orthogonality conditions (80) and the local conjugacy condition (82) as follows:

$$\int_0^1 B_*^T B_p dt + C_*^T C_p + (E_*^T E_p)_0 = 0, \qquad (89\text{-}1)$$

$$\int_0^1 B_*^T B_{*p} dt + C_*^T C_{*p} + (E_*^T E_{*p})_0 = 0, \qquad (89\text{-}2)$$

and

$$\int_0^1 M^T f_{ww} M_p dt + (M^T h_{ww} M_p)_0 + (M^T g_{ww} M_p)_1 = 0, \qquad (89\text{-}3)$$

where the subscript p denotes any conjugate gradient iteration preceding the present conjugate gradient iteration. While these equations do not guarantee convergence in a finite number of steps, they do guarantee that the algorithm generates a sequence of linearly independent vectors M(t), that is, a sequence of linearly independent variations per unit stepsize A(t), B(t), C.

7. Conjugate Gradient Phase: Practical Implementation

In this section, we summarize the results of Sections 5-6 and suggest practical ways of utilizing these results, within the following operating rules: (i) the use of second derivatives must be avoided; and (ii) the simultaneous determination of the directional coefficient γ and the stepsize α must be avoided. We refer to the general case where the functional (1) is nonquadratic and/or the constraints (2)-(5) are nonlinear.

7.1. Auxiliary Functions. The first step is to solve Eqs. (48)-(56) for a fictitious value of the directional coefficient, namely,

$$\gamma_* = 0 \,. \tag{90}$$

Using the solution technique of Section 4.4, we obtain the following auxiliary functions and multipliers:

$$A_*(t),\ B_*(t),\ C_*,\ \lambda_*(t),\ \sigma_*,\ \mu_* \,. \tag{91}$$

7.2. Directional Coefficient. The second step is to compute the actual value of the directional coefficient γ. For the first conjugate gradient phase, we set

$$\gamma = 0 \,. \tag{92}$$

For subsequent conjugate gradient phases, we set

$$\gamma = Q/\hat{Q}\,, \tag{93}$$

where

$$Q = \int_0^1 B_*^T B_* dt + C_*^T C_* + (E_*^T E_*)_0\,, \tag{94-1}$$

$$\hat{Q} = \int_0^1 \hat{B}_*^T \hat{B}_* dt + \hat{C}_*^T \hat{C}_* + (\hat{E}_*^T \hat{E}_*)_0\,. \tag{94-2}$$

In Eqs. (93)-(94), the symbols Q and $\hat{Q}$ denote the errors in the optimality conditions for the present conjugate gradient phase and the previous conjugate gradient phase, respectively.

The directional coefficient (93) is acceptable only if

$$\tilde{J}_\alpha(0) = -(Q + \gamma Z) < 0\,, \tag{95}$$

where Q is given by (94-1) and Z is given by

$$Z = \int_0^1 B_*^T \hat{B} dt + C_*^T \hat{C} + (E_*^T \hat{E})_0\,. \tag{96}$$

If Ineq. (95) is violated, then the directional coefficient (93) must be discarded and replaced by the value (92). This means that the algorithm must be restarted by replacing the conjugate gradient phase with an ordinary gradient phase.

7.3. Basic Functions. The third step is to compute the

basic functions A(t), B(t), C and the multipliers $\lambda(t)$, σ, μ. This is done with the following formulas:

$$A(t) = A_*(t) + \gamma\hat{A}(t), \qquad B(t) = B_*(t) + \gamma\hat{B}(t), \qquad C = C_* + \gamma\hat{C}, \tag{97}$$

$$\lambda(t) = \lambda_*(t), \qquad \sigma = \sigma_*, \qquad \mu = \mu_*. \tag{98}$$

Therefore, in the practical implementation of the algorithm, the basic functions A(t), B(t), C are computed using the formulas derived under the assumption of linear constraints.

7.4. <u>Stepsize</u>. With the basic functions (97) known, we consider the one-parameter family of varied functions

$$\tilde{x}(t) = x(t) + \alpha A(t), \qquad \tilde{u}(t) = u(t) + \alpha B(t), \qquad \tilde{\pi} = \pi + \alpha C . \tag{99}$$

After substitution of Eqs. (98)-(99) into (7) and (14), the following functions of the stepsize are obtained:

$$\tilde{J} = \tilde{J}(\alpha), \qquad \tilde{P} = \tilde{P}(\alpha). \tag{100}$$

Then, a one-dimensional search scheme is applied to (100-1), and a value of the stepsize α is selected for which the following relations are satisfied:

$$\tilde{J}(\alpha) < \tilde{J}(0), \qquad \tilde{P}(\alpha) \leq P_*, \qquad \tilde{\tau}(\alpha) > 0, \tag{101}$$

where τ is the final time and P_* is a preselected number, not necessarily small. Satisfaction of Ineq. (101-1) is possible

because of the descent property of the conjugate gradient phase. Ineq. (101-2) is introduced to prevent excessive constraint violation. And Ineq. (101-3) is required for problems with free final time.

Prior to the satisfaction of (101), a scanning process is employed, leading to the bracketing of the minimum point for $\tilde{J}(\alpha)$. This operation is then followed by a Hermitian cubic interpolation process (Ref. 9), which is stopped whenever the following relation is satisfied:[11]

$$|\tilde{J}_\alpha(\alpha)| \leq \varepsilon_3 \qquad \text{or} \qquad |\tilde{J}_\alpha(\alpha)/\tilde{J}_\alpha(0)| \leq \varepsilon_4, \tag{102}$$

subject to an upper limit for the number of search steps N_s. Once a stepsize α_o has been selected consistently with either (102) or the prescribed upper limit for the number of search steps, Ineqs. (101) must be checked. If satisfaction occurs, then the stepsize α_o is accepted. If any violation occurs, then the stepsize α_o must be bisected progressively until satisfaction of (101) is finally achieved.

[11] The symbols ε_3 and ε_4 denote small, preselected numbers.

8. Descent Property of a Cycle

A descent property exists for a complete conjugate gradient-restoration cycle under the assumption of small stepsizes. Let α_g denote the conjugate gradient stepsize and α_r the restoration stepsize. Simple manipulations, omitted for the sake of brevity, show that the conjugate gradient corrections are of $O(\alpha_g)$, while the restoration corrections are of $O(\alpha_r \alpha_g^2)$. Hence, for α_g sufficiently small, the restoration corrections are negligible with respect to the conjugate gradient corrections. Therefore, the restoration phase preserves the descent property of the conjugate gradient phase.

More specifically, let I_1, I_2, I_3 denote the values of the functional (1) at the beginning of the conjugate gradient phase, at the end of the conjugate gradient phase, and at the end of the subsequent restoration phase. Note that I_1 and I_2 are not comparable, since the constraints are not satisfied to the same accuracy. On the other hand, I_1 and I_3 are comparable, and the conjugate gradient stepsize α_g can be selected so that

$$I_3 < I_1 . \tag{103}$$

This inequality constitutes the descent property of a complete conjugate gradient-restoration cycle. In order to enforce it, one proceeds as follows. At the end of the restoration phase,

one must verify Ineq. (103). If it is satisfied, the next conjugate gradient phase is started; otherwise, the previous conjugate gradient stepsize is bisected as many times as needed until, after restoration, Ineq. (103) is satisfied.

9. Summary of the Algorithm

The sequential conjugate gradient-restoration algorithm solves optimal control problems involving a functional, subject to differential constraints and general boundary conditions. The algorithm is composed of a sequence of cycles, each cycle consisting of two phases, a conjugate gradient phase and a restoration phase. The objective of each cycle is to decrease the functional I, while the constraints are satisfied to the predetermined accuracy (17-1).

The decision parameters controlling the algorithm are the constraint error P and the optimality condition error Q [see Eqs. (14) and (15)]. If P violates Ineq. (17-1), the algorithm executes a restoration phase. If P satisfies Ineq. (17-1) and Q violates Ineq. (17-2), the algorithm executes a conjugate gradient phase. Finally, if P and Q satisfy Ineqs. (17), the algorithm stops: convergence has been achieved.

9.1. Restoration Phase. This phase involves one or more iterations and can be summarized as follows.

(a) Assume nominal functions $x(t)$, $u(t)$, π which satisfy condition (3), but violate at least one of conditions (2) and (4)-(5).

(b) For the nominal functions, solve the LTP-BVP (31)-(34) and (36)-(40) using the method of particular solutions. In this way, obtain the functions $A(t)$, $B(t)$, C and the

multipliers $\lambda(t)$, σ, μ.

(c) Using the functions in (b), compute the restoration stepsize by a one-dimensional search on the constraint error $\tilde{P}(\alpha)$. To this effect, perform a bisection process on α, starting from $\alpha_o = 1$, until Ineqs. (47) are satisfied.

(d) Once the restoration stepsize is known, compute the varied functions $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ with Eqs. (22).

(e) Verify whether the varied functions in (d) satisfy Ineq. (17-1). If this is the case, the restoration phase is terminated. Otherwise, return to (a) and continue the process until satisfaction of (17-1) occurs.

9.2. <u>Conjugate Gradient Phase</u>. This phase involves a single iteration and can be summarized as follows.

(a) Assume nominal functions $x(t)$, $u(t)$, π which satisfy the constraints (2)-(5) within the preselected accuracy (17-1).

(b) For the nominal functions and for $\gamma_* = 0$, solve the LTP-BVP (48)-(56) using the method of particular solutions. In this way, obtain the auxiliary functions $A_*(t)$, $B_*(t)$, C_* and the multipliers $\lambda_*(t)$, σ_*, μ_*.

(c) Set the directional coefficient γ at the level (92) for the first conjugate gradient phase and at the level (93) for any subsequent conjugate gradient phase. In the latter case, accept the directional coefficient only if Ineq. (95) is satisfied; otherwise, reset γ at the level (92).

(d) Compute the basic functions $A(t)$, $B(t)$, C and the multipliers $\lambda(t)$, σ, μ using Eqs. (97)-(98).

(e) Using the functions in (d), compute the conjugate gradient stepsize by a one-dimensional search on the augmented functional $\tilde{J}(\alpha)$ until satisfaction of Ineq. (102) occurs. Then, bisect the resulting stepsize α_o (if necessary), until satisfaction of Ineqs. (101) occurs.

(f) Once the conjugate gradient stepsize is known, compute the varied functions $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ with Eqs. (99).

9.3. <u>Conjugate Gradient-Restoration Cycle</u>. After the restoration phase is completed, verify whether Ineq. (103) is satisfied. If this is the case, start the next cycle of the sequential conjugate gradient-restoration algorithm. If not, return to the previous conjugate gradient phase and reduce the conjugate gradient stepsize (using a bisection process) until, after restoration, Ineq. (103) is satisfied.

10. Experimental Conditions

In order to evaluate the theory, nine examples were solved. The sequential conjugate gradient-restoration algorithm was programmed in FORTRAN IV, and the numerical results were obtained in double-precision arithmetic.

Computations were performed at Rice University using an IBM 370/155 computer. For each example, the interval of integration was divided into 100 steps. The differential equations were integrated using Hamming's modified predictor-corrector method with a special Runge-Kutta starting procedure (Ref. 10). The definite integrals I, J, P, Q were computed using a modified Simpson's rule. The method of particular solutions (Refs. 6-8) was used to solve the linear, two-point boundary-value problems associated with both the conjugate gradient phase and the restoration phase.

10.1. Convergence Conditions. The parameters ε_1, ε_2, ε_4 appearing in Ineqs. (17) and (102) were set at the levels[12]

$$\varepsilon_1 = E-08, \qquad \varepsilon_2 = E-04, \qquad \varepsilon_4 = E-03 . \tag{104}$$

The tolerance level (104-1) characterizes the restoration

---

[12] The symbol $E \pm ab$ stands for $10^{\pm ab}$.

phase; the tolerance levels (104-1) and (104-2), employed in combination, characterize the algorithm as a whole; and the tolerance level (104-3) characterizes the one-dimensional search for the conjugate gradient stepsize.

10.2. Safeguards. For the conjugate gradient phase, the parameter P_* appearing in Ineq. (101-2) was set at the level

$$P_* = 10. \qquad (105)$$

The tolerance level (105) limits the constraint violation which is permissible during the conjugate gradient phase. Also for the conjugate gradient phase, the number of Hermitian search steps required to satisfy Ineq. (102) was subject to the upper bound

$$N_s \leq 10 . \qquad (106)$$

10.3. Nonconvergence Conditions. The sequential conjugate gradient-restoration algorithm was programmed to stop whenever satisfaction of any of the following inequalities occurred:

$$\text{(i)} \quad N > 50 , \qquad (107)$$

$$\text{(ii)} \quad N_c > 30 , \qquad (108)$$

(iii) $N_r > 10$, (109)

(iv) $N_{br} > 10$, (110)

(v) $\tilde{J}_\alpha(0) \geq 0$, $\gamma = 0$, (111)

(vi) $N_{bg} > 10$, $\gamma = 0$, (112)

(vii) $N_{bc} > 10$, $\gamma = 0$. (113)

Here, N is the total number of iterations, N_c is the number of cycles, N_r is the number of restorative iterations per cycle, N_{br} is the number of bisections of the restoration stepsize required to satisfy Ineqs. (47), N_{bg} is the number of bisections of the conjugate gradient stepsize required to satisfy Ineqs. (101), N_{bc} is the number of bisections of the conjugate gradient stepsize required to satisfy Ineq. (103), and $\tilde{J}_\alpha(0)$ is the slope of the augmented functional at $\alpha = 0$.

Inequalities (107)-(108) apply to the algorithm as a whole. Satisfaction of (107) and/or (108) is indicative of extreme slowness of convergence.

Inequalities (109)-(110) apply to the restoration phase. Satisfaction of (109) is indicative of failure to produce a feasible solution in a reasonable number of restorative iterations. Satisfaction of (110) is indicative of extreme smallness of the restorative displacements.

Inequalities (111)-(112) apply to the conjugate gradient

phase. Satisfaction of (111) means that the descent property of the conjugate gradient phase does not hold, owing to numerical inaccuracy. Satisfaction of (112) is indicative of extreme smallness of the conjugate gradient displacements.

Inequality (113) applies to a complete conjugate gradient-restoration cycle. Satisfaction of (113) is indicative of extreme smallness of the displacements produced within a complete conjugate gradient-restoration cycle.

10.4. Restarting Conditions. The directional coefficient γ of the present conjugate gradient phase was reset at the level $\gamma = 0$ whenever satisfaction of any of the following inequalities occurred:

$$\text{(i)} \qquad \tilde{J}_{\alpha}(0) \geq 0 , \qquad \gamma = Q/\hat{Q} , \tag{114}$$

$$\text{(ii)} \qquad N_{bg} > 10 , \qquad \gamma = Q/\hat{Q} , \tag{115}$$

$$\text{(iii)} \qquad N_{bc} > 10 , \qquad \gamma = Q/\hat{Q} . \tag{116}$$

Satisfaction of (114) means that the descent property of the conjugate gradient phase does not hold. Satisfaction of (115) is indicative of extreme smallness of the conjugate gradient displacements. And satisfaction of (116) is indicative of extreme smallness of the displacements produced within a complete conjugate gradient-restoration cycle.

The directional coefficient γ of the next conjugate

gradient phase was reset at the level $\gamma = 0$ whenever satisfaction of any of the following inequalities occurred:

$$\text{(iv)} \quad 1 \leq N_{bg} \leq 10\,, \qquad \gamma = 0 \quad \text{or} \quad \gamma = Q/\hat{Q}\,, \tag{117}$$

$$\text{(v)} \quad 1 \leq N_{bc} \leq 10\,, \qquad \gamma = 0 \quad \text{or} \quad \gamma = Q/\hat{Q}\,. \tag{118}$$

Satisfaction of (117) or (118) is indicative of large violations of the orthogonality and conjugacy conditions, owing to the fact that the optimal conjugate gradient stepsize cannot be employed.

11. Numerical Examples

In this section, nine numerical examples are described employing scalar notation. In particular, the symbols $x_i(t)$, $i = 1,\dots,n$, denote the components of the state; the symbols $u_i(t)$, $i = 1,\dots,m$, denote the components of the control; and the symbols π_i, $i = 1,\dots,p$, denote the components of the parameter.

For all of the examples, a time normalization is used in order to simplify the numerical computations. Specifically, the actual time θ is replaced by the normalized time

$$t = \theta/\tau, \tag{119}$$

which is defined in such a way that $t = 0$ at the initial point and $t = 1$ at the final point. The actual final time τ, if it is free, is regarded as a component of the vector parameter π to be optimized. In this way, an optimal control problem with variable final time is converted into an optimal control problem with fixed final time.

Example 11.1. This is a linear-quadratic problem with (i) initial state partially given and (ii) fixed final time $\tau = 1$:

$$I = \int_0^1 (x_1^2 + x_2^2 + u_1^2)\,dt, \tag{120}$$

$$\dot{x}_1 = x_2 + u_1 , \qquad \dot{x}_2 = u_1 , \tag{121}$$

$$x_1(0) = 2 , \tag{122}$$

$$x_2(1) = 1 . \tag{123}$$

The assumed nominal functions are:

$$x_1(t) = 2 , \qquad x_2(t) = 1 , \qquad u_1(t) = -1 . \tag{124}$$

The numerical results are given in Tables 1-2. Convergence to the desired stopping condition occurs in $N = 4$ iterations, which include 1 restorative iteration and 3 conjugate gradient iterations.

Example 11.2. This is a linear-quadratic problem with (i) a linear relation between the components of the initial state and (ii) fixed final time $\tau = 1$:

$$I = \int_0^1 (x_1^2 + x_2^2 + u_1^2)\,dt , \tag{125}$$

$$\dot{x}_1 = x_2 + u_1 , \qquad \dot{x}_2 = u_1 , \tag{126}$$

$$x_1(0) + x_2(0) = 3 , \tag{127}$$

$$x_2(1) = 1 . \tag{128}$$

The assumed nominal functions are:

$$x_1(t) = 2, \qquad x_2(t) = 1, \qquad u_1(t) = -1. \tag{129}$$

The numerical results are given in Tables 3-4. Convergence to the desired stopping condition occurs in $N = 4$ iterations, which include 1 restorative iteration and 3 conjugate gradient iterations.

Example 11.3. This is a problem with (i) initial state given and (ii) fixed final time $\tau = 1$:

$$I = \int_0^1 (1 + x_1^2 + x_2^2 + u_1^2)\, dt\,, \tag{130}$$

$$\dot{x}_1 = u_1 - x_2^2\,, \qquad \dot{x}_2 = u_1 - x_1 x_2\,, \tag{131}$$

$$x_1(0) = 0, \qquad x_2(0) = 1\,, \tag{132}$$

$$x_1(1) = 1, \qquad x_2(1) = 2\,. \tag{133}$$

The assumed nominal functions are:

$$x_1(t) = t, \qquad x_2(t) = 1 + t, \qquad u_1(t) = 1\,. \tag{134}$$

The numerical results are given in Tables 5-6. Convergence to the desired stopping condition is achieved in $N = 7$ iterations, which include 5 restorative iterations and 2 conjugate

gradient iterations.

Example 11.4. This is a problem with (i) initial state given and (ii) fixed final time $\tau = 1$:

$$I = \int_0^1 (-2\cos u_1)\,dt\,, \tag{135}$$

$$\dot{x}_1 = 2\sin u_1 - 1, \qquad \dot{x}_2 = x_1\,, \tag{136}$$

$$x_1(0) = 0, \qquad x_2(0) = 0\,, \tag{137}$$

$$x_1(1) = 0, \qquad x_2(1) = 0.3\,. \tag{138}$$

The assumed nominal functions are :

$$x_1(t) = 0, \qquad x_2(t) = 0.3t\,, \qquad u_1(t) = 0\,. \tag{139}$$

The numerical results are given in Tables 7-8. Convergence to the desired stopping condition occurs in $N = 13$ iterations, which include 9 restorative iterations and 4 conjugate gradient iterations.

Example 11.5. This is a problem with (i) a nonlinear relation between the components of the initial state and (ii) fixed final time $\tau = 1$:

$$I = \int_0^1 (x_1^2 + x_2^2 + u_1^2)\,dt\,, \tag{140}$$

$$\dot{x}_1 = x_2 + u_1\,, \qquad \dot{x}_2 = u_1\,, \tag{141}$$

$$x_1^2(0) + x_2^2(0) = 5 , \tag{142}$$

$$x_2(1) = 1 . \tag{143}$$

The assumed nominal functions are:

$$x_1(t) = -2 , \qquad x_2(t) = 1 , \qquad u_1(t) = 0 . \tag{144}$$

The numerical results are given in Tables 9-10. Convergence to the desired stopping conditions occurs in $N = 7$ iterations, which include 4 restorative iterations and 3 conjugate gradient iterations.

<u>Example 11.6</u>. This is a minimum time problem with (i) a component of the initial state given, (ii) a nonlinear relation between the remaining components of the initial state, and (iii) free final time τ. After setting $\pi_1 = \tau$, the problem is as follows:

$$I = \pi_1 , \tag{145}$$

$$\dot{x}_1 = \pi_1 u_1, \qquad \dot{x}_2 = \pi_1 (x_1^2 - u_1^2), \qquad \dot{x}_3 = \pi_1 (u_1 - x_2^2 + x_1), \tag{146}$$

$$x_1(0) = 0, \qquad x_2^2(0) + x_3^2(0) = 1, \tag{147}$$

$$x_2(1) = 0, \qquad x_3(1) = 2. \tag{148}$$

The assumed nominal functions are:

$$x_1(t) = 0, \quad x_2(t) = 0, \quad x_3(t) = 1+t, \quad u_1(t) = 1, \quad \pi_1 = 1. \tag{149}$$

The numerical results are given in Tables 11-12. Convergence to the desired stopping condition occurs in $N = 9$ iterations, which include 7 restorative iterations and 2 conjugate gradient iterations.

Example 11.7. This is a minimum time problem with (i) a component of the initial state given, (ii) a nonlinear relation between the remaining components of the initial state, and (iii) free final time τ. After setting $\pi_1 = \tau$, the problem is as follows:

$$I = \pi_1, \tag{150}$$

$$\dot{x}_1 = \pi_1 x_3 \cos u_1, \quad \dot{x}_2 = \pi_1 x_3 \sin u_1, \quad \dot{x}_3 = \pi_1 \sin u_1, \tag{151}$$

$$x_1(0) = 0, \quad x_2(0) x_3(0) = 0, \tag{152}$$

$$x_1(1) = 1. \tag{153}$$

The assumed nominal functions are:

$$x_1(t) = t, \quad x_2(t) = 1, \quad x_3(t) = 0, \quad u_1(t) = 1, \quad \pi_1 = 1. \tag{154}$$

The numerical results are given in Tables 13-14. Convergence to the desired stopping condition occurs in $N = 13$ iterations, which include 10 restorative iterations and 3 conjugate gradient iterations.

Example 11.8. This is a minimum time problem with (i) a component of the initial state given, (ii) a nonlinear relation between the remaining components of the initial state, and (iii) free final time τ. After setting $\pi_1 = \tau$, the problem is as follows:

$$I = \pi_1, \tag{155}$$

$$\dot{x}_1 = \pi_1 x_3 \cos u_1, \qquad \dot{x}_2 = \pi_1 x_3 \sin u_1, \qquad \dot{x}_3 = \pi_1 \sin u_1, \tag{156}$$

$$x_1(0) = 0, \qquad x_2(0)x_3(0) = 0, \tag{157}$$

$$x_1^2(1) + x_3(1) = 1 . \tag{158}$$

The assumed nominal functions are:

$$x_1(t) = t, \quad x_2(t) = 1, \quad x_3(t) = 0, \quad u_1(t) = 1, \quad \pi_1 = 1. \tag{159}$$

The numerical results are given in Tables 15-16. Convergence to the desired stopping condition occurs in $N = 9$ iterations, which include 7 restorative iterations and 2 conjugate gradient iterations.

Example 11.9. This is a minimum time problem with (i) a component of the initial state given, (ii) a nonlinear relation between the remaining components of the initial state, and (iii) free final time τ. After setting $\pi_1 = \tau$, the

problem is as follows:

$$I = \pi_1 , \tag{160}$$

$$\dot{x}_1 = \pi_1 u_1, \qquad \dot{x}_2 = \pi_1 (x_1^2 - u_1^2), \qquad \dot{x}_3 = \pi_1 (u_1 - x_2^2 + x_1) , \tag{161}$$

$$x_1(0) = 0, \qquad x_2^2(0) + x_3^2(0) = 1 , \tag{162}$$

$$x_1(1) x_2(1) = 0, \qquad x_3(1) = 2. \tag{163}$$

The assumed nominal functions are:

$$x_1(t) = t, \quad x_2(t) = 0, \quad x_3(t) = 1 + t, \quad u_1(t) = 1, \quad \pi_1 = 1 . \tag{164}$$

The numerical results are given in Tables 17-18. Convergence to the desired stopping conditions occurs in $N = 9$ iterations, which include 7 restorative iterations and 2 conjugate gradient iterations.

Table 1. Convergence history, Example 11.1.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.10E+01		
1	0	1	1		0.22E-27	0.92E+01	7.72083
2	1	0	2	$\gamma=0$	0.16E-28	0.15E+00	5.89315
3	1	0	3	$\gamma\neq0$	0.40E-26	0.84E-03	5.83868
4	1	0	4	$\gamma\neq0$	0.77E-26	0.24E-06	5.83848

Table 2. Converged solution, Example 11.1.

t	x_1	x_2	u_1
0.0	2.0000	1.0198	-2.0336
0.1	1.9137	0.8410	-1.5512
0.2	1.8580	0.7081	-1.1115
0.3	1.8333	0.6175	-0.7059
0.4	1.8408	0.5661	-0.3264
0.5	1.8819	0.5516	0.0337
0.6	1.9586	0.5724	0.3815
0.7	2.0736	0.6277	0.7230
0.8	2.2299	0.7170	1.0646
0.9	2.4313	0.8408	1.4124
1.0	2.6822	1.0000	1.7729

$\tau = 1.00000$

Table 3. Convergence history, Example 11.2.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.10E+01		
1	0	1	1		0.13E-27	0.97E+01	5.99629
2	1	0	2	$\gamma=0$	0.61E-29	0.76E-01	4.23041
3	1	0	3	$\gamma\neq0$	0.40E-27	0.19E-03	4.20772
4	1	0	4	$\gamma\neq0$	0.14E-26	0.30E-07	4.20767

Table 4. Converged solution, Example 11.2.

t	x_1	x_2	u_1
0.0	1.2097	1.7902	-2.4220
0.1	1.1554	1.5684	-2.0209
0.2	1.1192	1.3848	-1.6564
0.3	1.1013	1.2361	-1.3225
0.4	1.1022	1.1195	-1.0137
0.5	1.1227	1.0327	-0.7250
0.6	1.1641	0.9740	-0.4516
0.7	1.2278	0.9420	-0.1891
0.8	1.3153	0.9359	0.0665
0.9	1.4290	0.9552	0.3198
1.0	1.5713	1.0000	0.5750

$\tau = 1.00000$

Table 5. Convergence history, Example 11.3.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.72E+01		
1	0	4	4		0.32E-10	0.97E+00	33.67701
2	1	1	6	$\gamma=0$	0.84E-13	0.50E-02	33.46606
3	1	0	7	$\gamma\neq0$	0.95E-08	0.14E-04	33.46465

Table 6. Converged solution, Example 11.3.

t	x_1	x_2	u_1
0.0	0.0000	1.0000	-8.3441
0.1	-0.7864	0.2775	-6.3703
0.2	-1.3015	-0.2371	-3.8640
0.3	-1.5841	-0.5630	-1.4827
0.4	-1.6737	-0.7172	0.4715
0.5	-1.6001	-0.7107	1.9964
0.6	-1.3776	-0.5433	3.2543
0.7	-1.0074	-0.2050	4.4921
0.8	-0.4872	0.3184	6.0506
0.9	0.1809	1.0418	8.4966
1.0	1.0000	2.0000	13.0501

$\tau = 1.00000$

Table 7. Convergence history, Example 11.4.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.10E+01		
1	0	4	4		0.17E-08	0.67E+00	-1.11665
2	1	2	7	$\gamma=0$	0.17E-11	0.34E-01	-1.16519
3	1	1	9	$\gamma\neq 0$	0.61E-09	0.31E-02	-1.16918
4	1	1	11	$\gamma\neq 0$	0.73E-12	0.37E-03	-1.16953
5	1	1	13	$\gamma\neq 0$	0.44E-15	0.47E-04	-1.16958

Table 8. Converged solution, Example 11.4.

t	x_1	x_2	u_1
0.0	0.0000	0.0000	1.3386
0.1	0.0937	0.0047	1.3031
0.2	0.1855	0.0186	1.2615
0.3	0.2741	0.0417	1.1995
0.4	0.3572	0.0733	1.1095
0.5	0.4304	0.1128	0.9775
0.6	0.4839	0.1587	0.7553
0.7	0.4924	0.2081	0.3688
0.8	0.4146	0.2543	-0.1523
0.9	0.2383	0.2877	-0.6112
1.0	0.0000	0.3000	-0.8962

$\tau = 1.00000$

Table 9. Convergence history, Example 11.5.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.10E+01		
1	0	1	1		0.00E-20	0.47E+00	3.33333
2	1	1	3	$\gamma=0$	0.10E-16	0.44E-02	3.24276
3	1	1	5	$\gamma\neq0$	0.35E-13	0.28E-02	3.24043
4	1	1	7	$\gamma\neq0$	0.84E-12	0.46E-05	3.23912

Table 10. Converged solution, Example 11.5.

t	x_1	x_2	u_1
0.0	-2.0314	0.9345	0.7621
0.1	-1.8703	0.9987	0.5292
0.2	-1.7250	1.0418	0.3391
0.3	-1.5935	1.0677	0.1851
0.4	-1.4739	1.0798	0.0617
0.5	-1.3647	1.0809	-0.0352
0.6	-1.2643	1.0736	-0.1089
0.7	-1.1712	1.0598	-0.1619
0.8	-1.0842	1.0418	-0.1958
0.9	-1.0015	1.0213	-0.2120
1.0	-0.9218	1.0000	-0.2114

$\tau = 1.00000$

Table 11. Convergence history, Example 11.6.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.20E+01		
1	0	3	3		0.75E-11	0.28E+00	1.01089
2	1	3	7	$\gamma=0$	0.51E-14	0.10E-02	0.86486
3	1	1	9	$\gamma\neq0$	0.15E-10	0.12E-04	0.86430

Table 12. Converged solution, Example 11.6.

t	x_1	x_2	x_3	u_1
0.0	0.0000	0.6465	0.7628	1.6143
0.1	0.1306	0.4493	0.8734	1.4180
0.2	0.2465	0.2970	0.9937	1.2700
0.3	0.3510	0.1783	1.1192	1.1527
0.4	0.4463	0.0871	1.2475	1.0539
0.5	0.5334	0.0199	1.3768	0.9643
0.6	0.6130	-0.0248	1.5059	0.8772
0.7	0.6850	-0.0483	1.6340	0.7876
0.8	0.7490	-0.0512	1.7598	0.6921
0.9	0.8044	-0.0345	1.8822	0.5890
1.0	0.8506	0.0000	2.0000	0.4782

$\tau = \pi_1 = 0.86430$

Table 13. Convergence history, Example 11.7.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.17E+01		
1	0	5	5		0.44E-16	0.25E+00	1.83370
2	1	2	8	$\gamma=0$	0.29E-09	0.42E-01	1.78266
3	1	2	11	$\gamma\neq0$	0.29E-13	0.16E-02	1.77286
4	1	1	13	$\gamma\neq0$	0.15E-11	0.24E-05	1.77245

Table 14. Converged solution, Example 11.7.

t	x_1	x_2	x_3	u_1
0.0	0.0000	1.0000	0.0000	1.5693
0.1	0.0016	1.0155	0.1765	1.4125
0.2	0.0129	1.0607	0.3486	1.2556
0.3	0.0426	1.1311	0.5121	1.0985
0.4	0.0974	1.2198	0.6630	0.9416
0.5	0.1819	1.3180	0.7975	0.7848
0.6	0.2974	1.4163	0.9125	0.6281
0.7	0.4426	1.5050	1.0050	0.4715
0.8	0.6129	1.5755	1.0728	0.3147
0.9	0.8016	1.6208	1.1143	0.1572
1.0	1.0000	1.6363	1.1280	-0.0014

$$\tau = \pi_1 = 1.77245$$

Table 15. Convergence history, Example 11.8.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.17E+01		
1	0	4	4		0.15E-13	0.45E-01	1.02941
2	1	2	7	$\gamma=0$	0.25E-17	0.32E-02	1.00403
3	1	1	9	$\gamma\neq 0$	0.14E-08	0.19E-04	1.00002

Table 16. Converged solution, Example 11.8.

t	x_1	x_2	x_3	u_1
0.0	0.0000	1.0000	0.0000	1.5648
0.1	0.0000	1.0050	0.1000	1.5639
0.2	0.0001	1.0200	0.1999	1.5630
0.3	0.0003	1.0450	0.2999	1.5623
0.4	0.0006	1.0800	0.3999	1.5615
0.5	0.0010	1.1250	0.4999	1.5608
0.6	0.0016	1.1800	0.5999	1.5600
0.7	0.0023	1.2450	0.6999	1.5590
0.8	0.0033	1.3200	0.7999	1.5579
0.9	0.0044	1.4050	0.8999	1.5567
1.0	0.0058	1.5000	0.9999	1.5551

$$\tau = \pi_1 = 1.00002$$

Table 17. Convergence history, Example 11.9.

N_c	N_g	N_r	N	γ	P	Q	I
0	0	0	0		0.86E+00		
1	0	3	3		0.53E-10	0.31E+00	1.02610
2	1	3	7	$\gamma=0$	0.24E-12	0.69E-03	0.86466
3	1	1	9	$\gamma\neq0$	0.33E-11	0.48E-05	0.86430

Table 18. Converged solution, Example 11.9.

t	x_1	x_2	x_3	u_1
0.0	0.0000	0.6492	0.7605	1.6273
0.1	0.1314	0.4496	0.8718	1.4247
0.2	0.2477	0.2961	0.9926	1.2734
0.3	0.3524	0.1770	1.1185	1.1545
0.4	0.4478	0.0856	1.2470	1.0547
0.5	0.5351	0.0185	1.3766	0.9645
0.6	0.6146	-0.0260	1.5059	0.8768
0.7	0.6866	-0.0492	1.6339	0.7868
0.8	0.7505	-0.0518	1.7598	0.6910
0.9	0.8058	-0.0349	1.8822	0.5876
1.0	0.8518	0.0000	2.0000	0.4763

$\tau = \pi_1 = 0.86430$

12. Discussion

The examples presented in Section 11 were solved with both the sequential conjugate gradient-restoration algorithm (SCGRA) of this report and the sequential ordinary gradient-restoration algorithm (SOGRA) of Ref. 5. This was done in order to gain perspective on the relative merit of SCGRA vis-a-vis SOGRA.

The comparative results[13] are presented in Tables 19-23, where the number of iterations N required to achieve different tolerance levels for the error in the optimality conditions Q is given for a fixed tolerance level of the constraint error $P \leq E\text{-}08$. Also shown in the tables are the values obtained for the objective functional I.

Cumulative results for the nine examples investigated are given in Table 24. Here, the total number of iterations for convergence ΣN is presented as a function of the tolerance level chosen for the error in the optimality conditions Q, for a fixed tolerance level in the constraint error $P \leq E\text{-}08$. In this comparative study, tolerance levels in the range $Q \leq E\text{-}02$ to $Q \leq E\text{-}06$ were chosen for the error in the optimality conditions.

---

[13] In Tables 19-23, the symbol LQ stands for a linear-quadratic problem, and the symbol NLQ stands for a nonlinear and/or nonquadratic problem.

From Tables 19-24, the following observations can be made.

(i) There is no saving in number of iterations for $Q \leq E\text{-}02$ and $Q \leq E\text{-}03$. The saving is 7% for $Q \leq E\text{-}04$, 14% for $Q \leq E\text{-}05$, and 18% for $Q \leq E\text{-}06$. Clearly, the relative advantage of SCGRA with respect to SOGRA increases by imposing a tighter tolerance level on the error in the optimality conditions.

(ii) On the average, the relative advantage of SCGRA with respect to SOGRA is greater for problems with free endpoints than for problems with fixed endpoints.

(iii) On the average, the relative advantage of SCGRA with respect to SOGRA is greater for linear-quadratic problems than for nonlinear and/or nonquadratic problems.

It must be noted that the experiments performed show that the computer time per iteration is roughly the same for SCGRA and SOGRA. Therefore, the conclusions pertaining to savings in number of iterations also apply to savings in computer time.

Table 19. Results for $P \leq E-08$ and $Q \leq E-02$.

Example	Type	SCGRA		SOGRA	
		N	I	N	I
11.1	LQ	3	5.83868	3	5.84026
11.2	LQ	3	4.20772	3	4.20801
11.3	NLQ	6	33.46606	6	33.46606
11.4	NLQ	9	-1.16918	9	-1.16923
11.5	NLQ	3	3.24276	3	3.24276
11.6	NLQ	7	0.86486	7	0.86486
11.7	NLQ	11	1.77286	10	1.77262
11.8	NLQ	7	1.00403	7	1.00403
11.9	NLQ	7	0.86466	7	0.86466

Table 20. Results for $P \leq E-08$ and $Q \leq E-03$.

Example	Type	SCGRA		SOGRA	
		N	I	N	I
11.1	LQ	3	5.83868	4	5.83854
11.2	LQ	3	4.20772	4	4.20768
11.3	NLQ	7	33.46465	7	33.46484
11.4	NLQ	11	-1.16953	11	-1.16950
11.5	NLQ	7	3.23912	7	3.23965
11.6	NLQ	9	0.86430	9	0.86431
11.7	NLQ	13	1.77245	10	1.77262
11.8	NLQ	9	1.00002	9	1.00064
11.9	NLQ	7	0.86466	7	0.86466

Table 21. Results for P $\leq$ E-08 and Q $\leq$ E-04.

Example	Type	SCGRA N	SCGRA I	SOGRA N	SOGRA I
11.1	LQ	4	5.83848	5	5.83848
11.2	LQ	4	4.20767	4	4.20768
11.3	NLQ	7	33.46465	7	33.46484
11.4	NLQ	13	-1.16958	13	-1.16964
11.5	NLQ	7	3.23912	11	3.23921
11.6	NLQ	9	0.86430	9	0.86431
11.7	NLQ	13	1.77245	12	1.77245
11.8	NLQ	9	1.00002	11	1.00012
11.9	NLQ	9	0.86430	9	0.86430

Table 22. Results for P $\leq$ E-08 and Q $\leq$ E-05.

Example	Type	SCGRA N	SCGRA I	SOGRA N	SOGRA I
11.1	LQ	4	5.83848	6	5.83848
11.2	LQ	4	4.20767	5	4.20767
11.3	NLQ	8	33.46464	8	33.46483
11.4	NLQ	14	-1.16962	15	-1.16965
11.5	NLQ	7	3.23912	14	3.23917
11.6	NLQ	10	0.86426	10	0.86429
11.7	NLQ	13	1.77245	12	1.77245
11.8	NLQ	11	1.00000	14	0.99998
11.9	NLQ	9	0.86430	9	0.86430

Table 23. Results for $P \leq E-08$ and $Q \leq E-06$.

Example	Type	SCGRA		SOGRA	
		N	I	N	I
11.1	LQ	4	5.83848	6	5.83848
11.2	LQ	4	4.20767	6	4.20767
11.3	NLQ	8	33.46464	8	33.46483
11.4	NLQ	15	-1.16962	17	-1.16965
11.5	NLQ	8	3.23912	17	3.23917
11.6	NLQ	10	0.86426	11	0.86429
11.7	NLQ	14	1.77245	13	1.77245
11.8	NLQ	12	0.99997	16	0.99997
11.9	NLQ	10	0.86428	10	0.86429

Table 24. Cumulative number of iterations for convergence, $P \leq E-08$.

Q	SCGRA	SOGRA
	ΣN	ΣN
$Q \leq E-02$	56	55
$Q \leq E-03$	69	68
$Q \leq E-04$	75	81
$Q \leq E-05$	80	93
$Q \leq E-06$	85	104

13. Conclusions

In this report, a new member of the family of sequential gradient-restoration algorithms for the solution of optimal control problems is presented. This is an algorithm of the conjugate gradient type and solves the problem represented by Eqs. (1)-(5): Minimize a functional subject to differential constraints and general boundary conditions.

The algorithm presented here differs from those of Refs. 3-4, in that it is not required that the state vector be given at the initial point. Instead, the initial conditions can be absolutely general. In analogy with Refs. 3-4, the present algorithm is capable of handling general final conditions; therefore, it is suitable for the solution of optimal control problems with general boundary conditions.

The importance of the present algorithm lies in that many optimal control problems either arise naturally in the present format or can be brought to such a format by means of suitable transformations (see Ref. 2). Therefore, a great variety of optimal control problems can be handled, as it is shown by the numerical examples presented.

Nine numerical examples are presented to illustrate the performance of the algorithm. The numerical results show the feasibility as well as the convergence characteristics of the algorithm. A comparative analysis of the sequential conjugate

gradient-restoration algorithm and the sequential ordinary gradient-restoration algorithm shows that the relative advantage of SCGRA with respect to SOGRA increases by imposing a tighter tolerance level on the error in the optimality conditions.

In summary, the new member of the family of sequential gradient-restoration algorithms described here has the following properties: (i) it retains the robustness, reliability, and convergence characterisitics of the algorithms discussed in Refs. 3-4; (ii) it is able to handle all of the optimal control problems treated in Refs. 3-4; and (iii) it has the additional capability of handling optimal control problems with general boundary conditions.

References

1. MIELE, A., PRITCHARD, R.E., and DAMOULAKIS, J.N., Sequential Gradient-Restoration Algorithm for Optimal Control Problems, Journal of Optimization Theory and Applications, Vol. 5, No. 4, 1970.

2. MIELE, A., Recent Advances in Gradient Algorithms for Optimal Control Problems, Journal of Optimization Theory and Applications, Vol. 17, Nos. 5-6, 1975.

3. HEIDEMAN, J.C., and LEVY, A.V., Sequential Conjugate Gradient-Restoration Algorithm for Optimal Control Problems, Part 1, Theory, Journal of Optimization Theory and Applications, Vol. 15, No. 2, 1975.

4. HEIDEMAN, J.C., and LEVY, A.V., Sequential Conjugate Gradient-Restoration Algorithm for Optimal Control Problems, Part 2, Examples, Journal of Optimization Theory and Applications, Vol. 15, No. 2, 1975.

5. GONZALEZ, S., and MIELE, A., Sequential Gradient-Restoration Algorithm for Optimal Control Problems with General Boundary Conditions, Rice University, Aero-Astronautics Report No. 142, 1978.

6. MIELE, A., Method of Particular Solutions for Linear, Two-Point Boundary-Value Problems, Journal of Optimization Theory and Applications, Vol. 2, No. 4, 1968.

7. MIELE, A., and IYER, R.R., General Technique for Solving Nonlinear, Two-Point Boundary-Value Problems via the Method of Particular Solutions, Journal of Optimization Theory and Applications, Vol. 5, No. 5, 1970.

8. MIELE, A., and IYER, R.R., Modified Quasilinearization Method for Solving Nonlinear, Two-Point Boundary-Value Problems, Journal of Mathematical Analysis and Applications, Vol. 36, No. 3, 1971.

9. MIELE, A., BONARDO, F., and GONZALEZ, S., Modifications and Alternatives to the Cubic Interpolation Process for One-Dimensional Search, Rice University, Aero-Astronautics Report No. 135, 1976.

10. RALSTON, A., Numerical Integration Methods for the Solution of Ordinary Differential Equations, Mathematical Methods for Digital Computers, Vol. 1, Edited by A. Ralston and H.S. Wilf, John Wiley and Sons, New York, New York, 1960.

Additional Bibliography

11. KNAPP, C.H., The Maximum Principle and the Method of Gradients, IEEE Transactions on Automatic Control, Vol. 11, No. 4, 1966.

12. LASTMAN, G.J., A Modified Newton's Method for Solving Trajectory Optimization Problems, AIAA Journal, Vol. 6 No. 5, 1968.

13. LASTMAN, G.J., and TAPLEY, B.D., Optimization of Non-linear Systems with Inequality Constraints Explicitly Containing the Control, International Journal of Control, Vol. 12, No. 3, 1970.

14. HONTOIR, Y., and CRUZ, J.B., JR., A Manifold Imbedding Algorithm for Optimization Problems, Automatica, Vol. 8, No. 5, 1972.